Dynamic *in Vivo* Binding of STAT5 to Growth Hormone-Regulated Genes in Intact Rat Liver. Sex-Specific Binding at Low- But Not High-Affinity STAT5 Sites

Ekaterina V. Laz, Aarathi Sugathan, and David J. Waxman

Division of Cell and Molecular Biology, Department of Biology, Boston University, Boston, Massachusetts 02215

Phylogenetic footprinting was used to predict functional transcription factor binding sites (TFBS) for signal transducer and activator of transcription (STAT) 5, a GH-activated transcription factor, in the GH-responsive genes *IGF-I*, *SOCS2*, and *HNF6*. Each gene, including upstream (100 kb) and downstream regions (25 kb), was aligned across four species and searched for conserved STAT5-binding sites using TFBS matrices. Predicted sites were classified as paired or single and whether or not they matched the STAT5 consensus sequence $TTCN_3GAA$. Fifty-seven of the predicted genomic regions were assayed by chromatin immunoprecipitation from male rat liver with high STAT5 activity. STAT5 binding was enriched (up to 24-fold) at eight genomic regions of *IGF-I*, including three novel regions in the second intron, and at four regions of *SOCS2*, including three novel upstream sites. STAT5 binding to *HNF6* was modestly enriched (up to 3-fold) at one consensus site and two novel, nonconsensus sites. Overall, 14 of 17 identified sites were paired STAT5 sites. STAT5 binding to these sites was dynamic in male rat liver, cycling on and off in response to each plasma GH pulse. Moreover, sex-specific STAT5 binding was apparent; in female rat liver, where nuclear STAT5 activity is generally low, STAT5 binding to *IGF-I* and *SOCS2* was limited to high-affinity sites. Analysis of the verified STAT5 binding sites indicated that STAT5 TFBS matrix 459 in combination with a STAT5 consensus sequence was the best predictor of STAT5 binding to these three genes. Using these criteria, multiple novel STAT5 binding sites were identified and then verified in several other GH-inducible genes, including *MUP* genes, where male-specific gene expression was associated with male-specific STAT5 binding to multiple low-affinity STAT5 sites. (***Molecular Endocrinology* 23: 1242–1254, 2009)**

GH, a pituitary-secreted polypeptide hormone, regulates a variety of metabolic processes, including fatty acid oxidation, amino acid uptake, and protein synthesis (1). The primary targets of GH include liver, muscle, and adipose tissue. GH is secreted from the pituitary in a sex-specific manner in rodents and humans (2–5). In adult male rats, peaks of GH secretion occur every 3–4 h and are separated by periods when GH is virtually undetectable (episodic GH profile), whereas in adult female rats, plasma GH peaks are more irregular and basal hormone levels are elevated compared with males (continuous GH profile) (2, 6). These sexually dimorphic plasma GH profiles establish and maintain sex differences in longitudinal bone growth as well as sex differences in the expression of a large number of genes in the liver (7–9).

GH binding to its cell surface receptor stimulates transphosphorylation of the GH receptor-associated Janus kinase 2 (JAK2) and activation of several downstream intracellular signaling pathways (10–13), including those mediated by signal transducer and activator of transcription (STAT) 5b. STAT5b is a key transcriptional regulator of GH signaling in the liver (14–18). STAT5b is activated by JAK2-dependent phosphorylation of tyrosine residue 699, which enables STAT5b to dimerize and translocate to the nucleus, where it binds DNA and activates transcription of target genes (19). DNA response elements for STAT5b and a closely related family member, STAT5a, collectively referred to as STAT5, can be represented by the consensus sequence $TTCN_3GAA$ (20). STAT5 is sensitive to the signal dynamics of GH stimulation (pulsatile *vs.* near continuous) (21–

ISSN Print 0888-8809 ISSN Online 1944-9917
Printed in U.S.A.

doi: 10.1210/me.2008-0449 Received December 2, 2008. Accepted April 20, 2009.
First Published Online May 7, 2009

Abbreviations: ChIP, Chromatin immunoprecipitation; HNF6, hepatocyte nuclear factor 6; JAK2, Janus kinase 2; NPV, negative predictive value; nt, nucleotide; PPV, positive predictive value; qPCR, quantitative real-time PCR; STAT, signal transducer and activator of transcription; SOCS2, suppressor of cytokine signaling 2; TFBS, transcription factor binding site.

23) and displays differential responsiveness to plasma GH stimulation in male and female rat liver (24). In male rats, the pool of liver STAT5 protein is repeatedly activated by each incoming plasma GH pulse; thus, there is a strong positive correlation between the plasma GH profile and the activity of STAT5 in the liver, with STAT5 activity levels being high during the upswing of a GH secretory episode and undetectable during the plasma GH trough periods (25, 26). Female rats have substantially lower liver STAT5 activity compared with male peak levels, but their basal (interpeak) STAT5 activity, although low, is measurably higher than the basal level in males (27).

IGF-I is a direct target of liver STAT5 (28–31). *IGF-I* mediates the effects of GH on somatic growth and tissue maintenance (1, 32, 33). The majority of circulating *IGF-I* is produced in the liver, where GH induces its expression (34). Using chromatin immunoprecipitation (ChIP), two DNA regions, each containing a pair of STAT5 binding sites, have been identified in the rat *IGF-I* locus: region RE-1/RE-2, located 75 kb upstream of the transcription start site, and region GHRE-1/GHRE-2, located in the first intron (28, 31). STAT5 rapidly binds to these two regions in liver after treatment of hypophysectomized rats with a supraphysiological dose of GH, which induces transcription of the *IGF-I* gene (28, 31). The two 5′ distal STAT5 binding sites (RE-1/RE-2) were also identified in the human *IGF-I* gene through mapping of STAT5-binding enhancers (35). In addition, three novel *IGF-I* regions containing a total of five consensus STAT5 binding sites were recently identified by ChIP analysis of GH-treated mouse liver (29).

STAT5 binding sites have been identified in liver for a limited number of other GH-inducible genes. Suppressor of cytokine signaling 2 (*SOCS2*) encodes a GH-inducible negative regulator of JAK/STAT signaling that acts on GH receptor and other cytokine receptor signaling pathways (36, 37). A GH-response element containing a pair of STAT5 binding sites that is conserved between rat and human *SOCS2* was identified in the rat (38). As shown by ChIP analysis, GH rapidly stimulates binding of STAT5 to this region in the liver of hypophysectomized rats, coincident with the induction of *SOCS2* gene transcription (38). Hepatocyte nuclear factor 6 (*HNF6*) may also be a direct target of GH-activated STAT5 (39). *HNF6* is a female-predominant (40, 41), liver-enriched transcription factor that regulates the transcription of a variety of genes (42–44), including certain sex-dependent, GH-responsive CYP genes (40, 45, 46). A consensus STAT5 binding site in the *HNF6* promoter was shown to bind STAT5 in a GH-dependent manner *in vitro*, and this STAT5 binding site was required for GH-stimulated transcription of an *HNF6* promoter-reporter gene (39).

The known STAT5 binding sites in *IGF-I*, *SOCS2*, and *HNF6* were all identified under nonphysiological conditions, *i.e.* in livers of hypophysectomized rats or mice given a supraphysiological dose of GH or *in vitro*. Conceivably, additional binding sites might be bound by STAT5 in liver *in vivo*, where STAT5 binding to chromatin may vary during the course of a naturally occurring plasma GH pulse in response to changes in intranuclear concentrations of tyrosine-phosphorylated STAT5. Furthermore, the extent to which STAT5 sites are occupied might differ between males and females, due to the sex differences in plasma GH profiles. These and related issues are examined in the present study, where we use phylogenetic footprinting to predict STAT5 binding sites in the *IGF-I*, *SOCS2*, and *HNF6* genes, and we test these sites experimentally by ChIP analysis using livers of intact male and female rats. On the basis of our findings, a refined computational approach is used to predict, and then verify, liver STAT5 binding to several other GH target genes, including *MUP* genes (47), which are expressed in a STAT5b-dependent and male-specific manner subject to GH regulation.

Results

Prediction and experimental evaluation of STAT5 binding sites for *IGF-I*, *SOCS2*, and *HNF6*

Table 1 presents a summary of STAT5 binding sites predicted in rat and three other species using a set of nine STAT5 transcription factor binding sites (TFBS) matrices (supplemental Table S1, published as supplemental data on The Endocrine Society's Journals Online web site at http://mend.endojournals.org). A majority (88%) of the predicted STAT5 sites fall into the paired zero consensus and single nonconsensus categories, *i.e.*

TABLE 1. Summary of predicted STAT5 binding sites

	Predictions			
Type	All sites[a]	Sites found in all four species[b]	Tested[c]	Verified[c]
IGF-I				
Paired 2 consensus	2	2	2	2
Paired 1 consensus	32	12	10	5
Paired 0 consensus	207	25	11	1
Single consensus	18	4	3	2
Single nonconsensus	115	16	0	0
Total	374	59	26	10
SOCS2				
Paired 2 consensus	2	1	2	2
Paired 1 consensus	11	1	1	1
Paired 0 consensus	135	13	8	0
Single consensus	12	2	3	1
Single nonconsensus	80	3	2	0
Total	240	20	16	4
HNF6				
Paired 2 consensus	1	0	1	0
Paired 1 consensus	17	3	4	1
Paired 0 consensus	163	16	16	2
Single consensus	9	1	1	0
Single nonconsensus	91	9	0	0
Total	281	29	22	3

For experimental validation, preference was given to the sites conserved in all four species (rat, mouse, human, and dog). Adjacent sites located in the proximity of amplicons for other sites were in many cases not considered if not found in all four species.

[a] STAT5 sites predicted for the rat genes based on a set of nine TFBS matrices (see supplemental Table S1).

[b] Predicted rat STAT5 sites that are also present in mouse, human, and dog.

[c] Testing and verification is based on results for 57 amplicons (64 sites) presented in Fig. 1 and data in Fig. 4.

they do not contain the STAT5 binding site consensus sequence $TTCN_3GAA$. To test these STAT5 binding site predictions, we primarily considered sites found in all four species because DNA regulatory elements are often conserved evolutionarily (48, 49). Predicted STAT5-binding sites were tested by ChIP followed by qPCR (supplemental Table S2) using rat liver chromatin prepared from an untreated male with high liver STAT5 activity (see Fig. 2, sample 11). A total of 26, 16, and 22 predicted STAT5 sites were tested in the *IGF-I*, *SOCS2*, and *HNF6* genes, respectively. Predicted STAT5-binding sites in close proximity to each other [within 150 nucleotides (nt)] were considered together as part of a single STAT5 binding region. The chromosomal locations of the tested STAT5 binding sites and the corresponding amplicons used for experimental validation are shown in supplemental Table S3. The abundance of STAT5 binding in the immunoprecipitated samples was normalized to DNA input and compared with the abundance of signal for the negative control, located within a 2300-nt segment in the 5′ distal region of rat *IGF-I* that is devoid of any predicted STAT5 binding sites. Eight regions in the *IGF-I* gene (corresponding to a total of 10 STAT5 sites) were enriched at least 2-fold over the negative control in STAT5 antibody precipitates (Fig. 1A; sites numbered in *color*). The highest enrichment (24-fold) was observed for region 296, which contains three consensus STAT5 binding sites. Three of the eight regions are novel STAT5 binding regions, all located within intron 2 (regions 217, 232, and 260) (Table 2).

In the case of *SOCS2*, five regions were enriched more than 2-fold compared with the negative control (Fig. 1B). All five regions (193, 199, 221/222, 224, and 225) are located 5′ to the rat *SOCS2* gene. The weak ChIP signal of *SOCS2* site 225 likely comes from the adjacent strong STAT5 site 224, insofar as site 225 does not bind STAT5 *in vitro* (see below). Site 224 was previously shown to bind STAT5 in the livers of hypophysectomized rats treated with GH (38), whereas the other three sites represent novel STAT5 binding regions. In the case of *HNF6*, three STAT5-binding regions exhibited a 1.9- to 3-fold enrichment over the negative control (Fig. 1C). *HNF6* site 181 was previously shown to bind STAT5 in a GH-dependent manner by EMSA analysis (39), whereas the other two *HNF6* sites, 148 and 157, are novel and contain paired nonconsensus STAT5 sequences (Table 2). The enrichment of STAT5 binding to these regions of *HNF6* is considerably lower than that for the strongest STAT5 binding regions in the *IGF-I* and *SOCS2* genes (Fig. 1). A fourth *HNF6* region, 234, was approximately 2-fold enriched in STAT5 binding with respect to the IgG control but was only 1.4-fold enriched relative to the negative control and was not considered further.

Sex dependence of STAT5 binding *in vivo*

Liver samples from individual untreated male and female rats were used in STAT5 ChIP analysis to investigate the relationship between the temporal pattern of liver STAT5 activation (25, 26) and STAT5 binding to chromatin. The STAT5 activity status of each liver is shown in Fig. 2. STAT5 bands were quantified and expressed as a percentage of the STAT5 signal in liver 11 (set at 100%). We assayed four males with high STAT5 activity (Fig. 2, lanes 11–14, 60–100% of sample 11), four males with intermediate STAT5 activity (lanes 7–10, 3–34%), two males with no detectable STAT5 activity (lane 6 and another, similar liver sample, 0.2–0.3%). Four female livers, all having STAT5 activity that was very low but detectable were also assayed (lanes 1–4, 0.8–1.4% of sample 11; see intensified image at *bottom* for detection of the low female liver signals). Samples were subject to STAT5 ChIP analysis, with real-time PCR quantification of each of the genomic regions enriched for STAT5 binding in Fig. 1. Clear differences in STAT5 binding were observed when comparing male liver samples with high *vs.* intermediate *vs.* very low or undetectable STAT5 activity. The majority of STAT5 binding regions were occupied in male samples with high STAT5 activity (Fig. 3). These same regions exhibited lower STAT5 binding in the intermediate STAT5 activity male group and no STAT5 binding, compared with the negative control, in the STAT5 activity-deficient male group. Because there is a strong positive correlation between the plasma GH profile and liver STAT5 activity (25, 26), these results indicate that STAT5 cycles off of its chromatin binding sites during the plasma GH interpulse period. STAT5 binding at *IGF-I* region 260 and *HNF6* regions 148 and 157, detected in the initial screen (Fig. 1C), was not observed in Fig. 3 (<2-fold enrichment relative to the negative control), partially due to a lower sensitivity of the latter ChIP experiment, where the cross-linked chromatin was not purified and where 2- to 3-fold less DNA per ChIP sample was used compared with the initial screening. STAT5 binding in female liver samples, where STAT5 activity was marginally higher than that of the STAT5-negative males (Fig. 2, lanes 1–4 *vs.* lanes 5 and 6) was indistinguishable from that of the STAT5-negative males, except at *IGF-I* sites 196 and 304 and *SOCS2* sites 224 (and 225) (Fig. 3). *HNF6* bound STAT5 poorly compared with the STAT5 binding sites in *IGF-I* and *SOCS2*, in agreement with Fig. 1. Weak binding was observed at *HNF6* region 181 in males with high STAT5 activity levels but not in females or in other males (Fig. 3B).

In vitro STAT5 binding to predicted sites

Although the above ChIP assays allowed us to identify *in vivo* STAT5 binding regions, they have limited resolution and do not establish whether a particular genomic sequence is capable of STAT5 binding. We therefore performed a competitive EMSA assay to determine the intrinsic STAT5 binding activity of the STAT5 sites enriched in the ChIP assay. As shown in Fig. 4 and summarized in Table 2, EMSA probes corresponding to *IGF-I* sites 217, 296, 260, and 304; *SOCS2* sites 193M, 199, 221 and 224; and *HNF6* sites 148 and 181 all competed efficiently for STAT5 binding. Less extensive competition was seen with *IGF-I* sites 232 and 263 and *HNF6* site 157. These findings are consistent with these specific STAT5-binding sequences being responsible for the positive signals seen in the ChIP assay (Fig. 1). *IGF-I* site 232 and *HNF6* sites 148 and 157 correspond to nonconsensus STAT5 binding sequences. Probes representing *SOCS2* sites 222 and 225 and the negative control probe Oct-1 did not compete for STAT5 binding. The inability of SOCS2 site 225 to compete for STAT5 binding supports our suggestion, above, that its enrichment in ChIP reflects STAT5 binding to

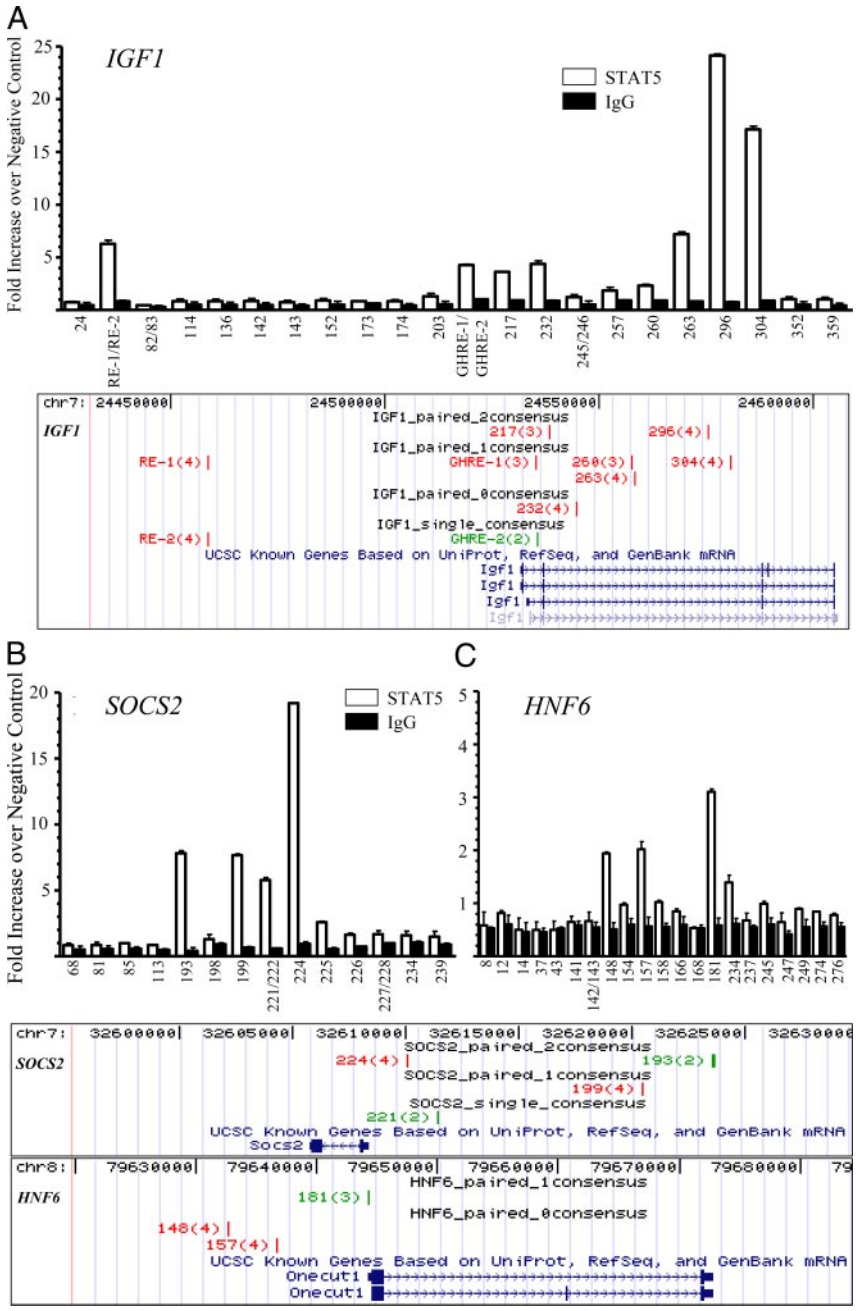

FIG. 1. STAT5 binding regions identified by ChIP. Chromatin samples prepared from untreated male liver with a high content of active STAT5 (sample 11 in Fig. 2) (35–60 μg DNA per sample) were precipitated with STAT5 antibody or normal rabbit IgG, and the abundance of STAT5 binding regions predicted in the *IGF-I* (*A*), *SOCS2* (*B*), or *HNF6* (*C*) loci was quantified by real-time PCR. Data were normalized to input DNA and expressed for each region as fold increase over the negative control (region within a 2300-bp segment in the 5′ distal region of rat *IGF-I*, which is devoid of any predicted STAT5 binding sites). Data are mean ± range values for two independent determinations. qPCR primers used to assay each site are shown in supplemental Table S2, and the chromosomal coordinates of the corresponding amplicons and predicted STAT5 sites are shown in supplemental Table S3. In cases where predicted STAT5 sites were in close proximity, a single PCR amplicon was used to interrogate both sites, as indicated (*e.g. IGF-I* sites 245 and 246, *SOCS2* sites 221 and 222, *etc.*). Genomic locations of the sites positive for STAT5 binding are shown *below* each graph, as visualized on the UCSC genome browser (http://genome.ucsc.edu). *Red* indicates sites present in all four species; *green* indicates sites present in three of the four species considered. *Numbers in parentheses* indicate the number of species (of four total) in which the local genomic region encompassing the site is at least 70% identical. All STAT5 sites shown on the genome browser window except for *IGF-I* site GHRE-2, *SOCS2* sites 193 and 221, and *HNF6* site 181 are present in all four species (rat, mouse, human, and dog).

the adjacent site 224. Similarly, STAT5 binding to *SOCS2* region 221/222 probably reflects STAT5 binding to site 221, a strong binding site, and not to site 222.

Scatchard analysis of STAT5 binding affinity

We investigated the possibility that the sites showing substantial STAT5 binding in female liver (*IGF-I* sites 296 and 304 and *SOCS2* site 224; Fig. 3) have high affinity for STAT5, which would enable them to bind STAT5, both in males during a plasma GH pulse, when STAT5 activity is high, and in females, where STAT5 activity is low but more persistent. EMSA assays revealed that these three sites were the best competitors for STAT5 binding among the probes tested (Fig. 5A). Scatchard plot analysis revealed that all three sites are high-affinity STAT5 binding sites, with dissociation constant (K_d) values of approximately 1–2 nM (Fig. 5B). In contrast, *SOCS2* site 199, which shows high STAT5 binding in high STAT5 activity males but low STAT5 binding in females (Fig. 3B), is a low-affinity binding site, K_d = 9.3 ± 0.7 nM (Fig. 5B), in agreement with the rank order of STAT5 binding in female liver *in vivo*.

Evaluation of STAT5 binding site prediction

Overall, 17 of the 64 STAT5 sites tested (27%) bound STAT5 in STAT5-positive males (Fig. 1 and Table 1). In an effort to improve the validation rate, we investigated which features distinguish predicted sites that bind STAT5 from those that do not. This evaluation was based on the 64 STAT5 binding sites indicated in Fig. 1 plus 15 additional sites that were conserved in fewer than four species but were located in close proximity to the 64 sites and therefore were indistinguishable from them by ChIP. These 79 sites were classified according to 19 possible predictors of STAT5 binding (see *Materials and Methods*); these include the presence of a paired STAT5 site, a STAT5 consensus sequence in rat or in other species, a STAT5 site predicted in one or more other species, and a STAT5 site that is matched by any of the nine available STAT5 matrices. Measures of accuracy of prediction, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy, were calculated for each predictor (supplemental Table S5). The predictors with the best combination of sensitivity and specificity were found to be 1) the presence of at least one consensus site in rat, and 2) a match with matrix 459. These two predictors also had the best accuracies and the best combinations of PPV and NPV (Table 3). This finding is

TABLE 2. Summary of STAT5 binding sites for rat genes enriched in ChIP assays

Site	Classification[a]	Sequence,[b] 5′–3′	Chromosomal location[a]	Verified by EMSA
IGF-I on Chr7(+)				
RE-1[c]	Paired 1 consensus	tctgtgttagtcaggaaaaTTCTAAGAAactgcctccagagagagg	24,458,482–24,458,527	Yes (Ref. 28)
RE-2[c]	Single consensus	tttTTCTTAGAAgta	24,458,733–24,458,747	Yes (Ref. 28)
GHRE-1[c]	Paired 1 consensus	ccgctcaccttgggggccTTCCTGGAAgaa	24,535,315–24,535,344	Yes
GHRE-2[c]	Single consensus	tgcTTCTTAGAAtga	24,535,399–24,535,413	Yes (Ref. 31)
217	Paired 2 consensus	catTTCTTTGAAgtgcaaggagTTCCTGGAAcct	24,538,140–24,538,173	Yes
232	Paired 0 consensus	ggatcccaagaaaaaccccttcccttgc	24,544,749–24,544,775	Yes
260	Paired 1 consensus	cattttaaacgtaagTTCTGAGAActg	24,557,417–24,557,443	Yes
263[c]	Paired 1 consensus	tctTTCAGGGAAatctaggaatatcagaaa	24,558,319–24,558,348	Yes
296[c]	Paired 2 consensus	ggcaactgtgaataagtttTTCGAAGAA ttg(6)gacttctgaggcaacggtctcca gTTCTCAGAAaggaaaTTCGCAGAAgtg	24,575,406–24,575,493	Yes
304[c]	Paired 1 consensus	tgaTTCCTAGAAaagatgacctcacccaac	24,580,690–24,580,719	Yes
SOCS2 on Chr7(−)				
193	Paired 2 consensus	atattattggaaatc(4)ctctgacaagca ctgtactaggaa(29)ttgTTCTTGGAA tgt(18)tgcTTCTCTGAAgttcaggtgc tcggtctacaaaatgtgatctatgtggaaag	32,623,539–32,623,677	Yes
199	Paired 1 consensus	tgcTTCTCAGAAtccgatgactaagccaggaatag	32,620,436–32,620,470	Yes
221	Single consensus	agaTTCCAAGAAaac	32,611,389–32,611,403	Yes
222	Paired 0 consensus	tagaattttctaaagagaaaaaaattactgcggataa	32,611,262–32,611,298	No
224[c]	Paired 2 consensus	gcggtcacgtgaggcggaTTCCTGGAA agTTCCTGGAAag cggcctccgcagcggc	32,610,008–32,610,063	Yes
225	Paired 0 consensus	tccttctcggcgtcgggaaatcttcggagcac (16)ctgttatccaaatttataatcctaataacct	32,609,489–32,609,567	No
HNF6 on Chr8(+)				
148	Paired 0 consensus	tgatacccagaattctattgaccatgg	79,632,576–79,632,602	Yes
157	Paired 0 consensus	tttccatcattaatgtcattactacgaacta(10) tgtcgttgggagccgagtttcacggtattg	79,636,722–79,636,792	Yes
181[c]	Paired 1 consensus	gagccgggggcagcaggaTTCTAAGAAaga	79,644,216–79,644,245	Yes
RGS3 on Chr5(+)				
113	Paired 1 consensus	ttgtgtgctcagaccataTTCTCAGAAtaa	79,706,585–79,706,614	ND
122	Paired 1 consensus	ggcctcatggcctcctatttgtggaagcaca ggatagtgactTTCCAAGAActgctctttgtttctctca	79,724,023–79,724,092	ND
SPIN2[d] on Chr6				
12 (2c)	Paired 1 consensus	gatTTCTGGGAAcatggactcatagtccct	128,461,625–128,461,654	ND
21 (2b)	Paired 0 consensus	attgtcccagaaatccacttcctctcagatcctcagaaatg	128,378,512–128,378,552	ND
26 (2a)	Paired 1 consensus	tgatTTCTCAGAAcatggattagtagaagcg	128,436,433–128,436,463	ND
54 (2b)	Paired 1 consensus	cgcttctactaatccatgTTCTGAGAAatca	128,386,919–128,386,949	ND
MUPs on Chr5(−)				
8 (*OBP3*)	Single consensus	gtcTTCTGAGAAtcc	78,179,829–78,179,843	Yes
9 (*OBP3*)	Single consensus	caaTTCATGGAAatt	78,179,664–78,179,678	Yes
29	Single consensus	Various	Various (see supplemental Table S3)	Yes
50	Single consensus	gtcTTCTGAGAAtcc	Various (see supplemental Table S3)	Yes

Chr, Chromosome. ND, not determined.

[a] Classification and chromosomal location of STAT5 binding sites are based on predictions made with all nine STAT5 TFBS matrices. The nt numbering is based on rat genome assembly rn4.

[b] Consensus STAT5 binding sequence (TTCNNNGAA) within a predicted STAT5 binding site is shown in *uppercase letters*. *Numbers in parentheses* indicate the nucleotide length of the intervening sequence that is not recognized by a STAT5 binding matrix.

[c] STAT5 binding sites identified in the present study that were also previously identified (Refs. 28, 29, 31, 35, 38, and 39).

[d] *Number in parentheses* indicates the *SPIN2* gene located closest to the STAT5 binding site shown.

consistent with the fact that matrix 459 (supplemental Table S1) is the only one of the nine STAT5 matrices that specifically describes a binding site for STAT5b, the major liver STAT5 form (50, 51). When the presence of a consensus STAT5 sequence and matrix 459 were combined as a single predictor, the specificity, accuracy, and PPV all increased (Table 3). Applying the combined predictor to the same three genes for sites conserved in at least three species resulted in correct predictions for eight of 10 *IGF-I* sites, two of the three *SOCS2* sites, but not the one *HNF6* site, with an overall validation rate of 71%, as compared with about 27% in the original analysis using all nine STAT5 matrices. Decreasing the minimum number of species sharing a site to two resulted in the correct prediction of all 14 validated consensus STAT5 sites and an overall validation rate of 61% (supplemental Table S6). These two condi-

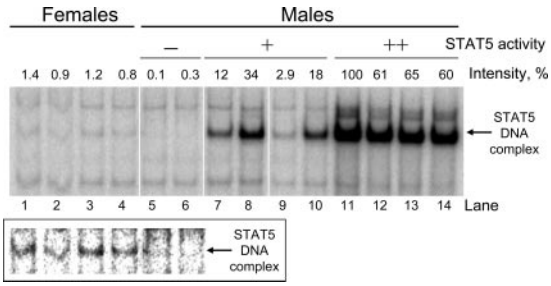

FIG. 2. EMSA analysis of STAT5 DNA-binding activity in individual adult rat livers. Homogenates prepared from individual rat livers were assayed for STAT5 binding using a ^{32}P-labeled STAT5 binding probe (STAT5 response element of the rat β-casein promoter; supplemental Table S4). Male livers were classified as high (++) (lanes 11–14), intermediate (+) (lanes 7–10), and no STAT5 activity (−) (lanes 5 and 6). Male liver 11 was used in ChIP experiments presented in Fig. 1 and was included with the other livers in ChIP experiments presented in Fig. 3. EMSA band intensities expressed as a percentage of liver 11 are shown *above* each lane. All 14 lanes were from the same gel and exposure. *Lower panel* shows lanes 1–6 at a higher intensity to better visualize the differences in STAT5 activity content between female samples (first four lanes) and STAT5 (−) male liver samples (last two lanes). All samples shown, except for liver 5, were used for the ChIP analyses shown in Fig. 3.

tions can therefore be used to shortlist the predicted STAT5 candidate sites for experimental validation.

Genes that do not exhibit a pattern of expression similar to that of *IGF-I* and *SOCS2* may contain STAT5 binding sites that do not match the STAT5 consensus sequence. Because matrix 459 is one of the best individual predictors (supplemental Table S5 and Fig. S1), it can be applied on its own to predict nonconsensus as well as consensus sites. However, this would bias results toward sites that resemble matrix 459, and it might be more desirable to predict nonconsensus binding sites that are closer to one of the other STAT5 matrices but with a more stringent TFBS matrix matching (Possum) score. A Possum score of at least 7, instead of the default value of 5, is not one of the best predictors when applied alone (supplemental Table S5); however, when applied as a second filter after selecting sites that are found in all four species, a Possum score of 7 gives results similar to those obtained using matrix 459; both predictors miss one verified site each, although the former predicts more unverified sites (supplemental Table S7).

Identification of STAT5 binding sites in other GH-responsive genes

Two of the conditions considered above, presence of a STAT5 consensus site and a match with matrix 459, were used to predict STAT5 binding sites in a set of 12 early GH-responsive genes, previously identified in rat liver by microarray analysis (*CALD1*, *GADD45G*, *NREP*, *SULT2A1*, *RGS3*, *SPIN2A*, *SPIN2B*, and five members of the *MUP* gene family) (52). These genes are all down-regulated in hypophysectomized compared with intact male rat liver and are rapidly induced (within 30–90 min) by a single, physiological GH injection, making them candidates for direct STAT5 target genes (supplemental Table S8). ChIP analysis revealed strong STAT5 binding to two sites in RGS3 (sites 113 and 122) and to four sites in the *SPIN2* genes (Fig. 6A, sites 12, 21, 26, and 54) (Table 2). Multiple functional STAT5 binding sites were found for the *MUP* genes (sites 8/9, site 29, and site 50, Fig. 6B). The amplicons representing sites 29 and 50 each mapped to six distinct genomic sequences associated with the eight known *MUP* genes (Fig. 6B) and thus could not be assigned to a specific *MUP* gene; most likely, multiple *MUP* region genomic positions contribute to the observed site 29 and site 50 ChIP signals. In the case of *MUP* sites 8/9, which are conserved in at least three *MUP* genes (*OBP3*, *MUP4*, and LOC259246), a gene-specific amplicon [designated 8/9 (OBP3)] could be designed. STAT5 binding to all of the active *MUP* binding regions was male specific in liver chromatin (Fig. 6C), consistent with the strong male specificity of *MUP* gene expression in rat liver (47, 52, 53). Scatchard plot analysis for binding of STAT5 to *OBP3* sites 8 and 9 (whose local sequences are also shared by sites 29 and 50 of *OBP3*, LOC259246, and several other *MUP* genes; supplemental Table S9) yielded K_d values of approximately 10–12 nM (Fig. 6D). The low affinity of these male-specific STAT5 sites further supports the hypothesis that STAT5 binding in females primarily occurs at high-affinity sites. None of the predicted STAT5 sites tested for *CALD1* (four sites), *GADD45G* (two sites), *NREP* (one site), or *SULT2A1* (one site) bound STAT5 in liver chromatin (data not shown).

Discussion

STAT5b is an essential mediator of GH action in the liver, where it regulates the transcription of many genes either directly or indirectly. In rat liver, at least 20% of the genes acutely stimulated by GH and a majority of the genes acutely suppressed by GH are dependent on STAT5b for their regulation (38, 54). This study investigated the utility of phylogenetic footprinting for discovery of STAT5 binding sites in GH-responsive genes. In addition, the dynamic effect of plasma GH profiles on STAT5 binding to chromatin was investigated in intact male and female rat liver. We used a computational approach to predict STAT5 binding sites, where sequences associated with each gene and its flanking DNA (100 kb upstream sequence and up to 25 kb downstream sequence) were first scanned using TFBS matrices for STAT5. Long genomic sequences were analyzed because of the emerging evidence that STAT5 often binds to genes outside of the traditional promoter region (28, 29, 55). Unlike previous studies, where only the consensus sequence $TTCN_3GAA$ was considered in predicting STAT5 binding sites (29, 38, 55), our approach used TFBS matrices, including matrices based on nonconsensus STAT5 sequences (56). The relevance of nonconsensus sequences is supported by a recent identification of a STAT5 binding region containing multiple nonconsensus sequences in the promoter of *C3ar1* (57). Because regulatory elements are often conserved between species (48, 49), predicted rat STAT5 sites that were conserved across the other three species examined (mouse, human, and dog) were given preference for experimental verification. Moreover, in contrast to previous studies, where STAT5 binding sites associated with *IGF-I*, *SOCS2*, and *HNF6* were identified either under nonphysiological conditions or *in vitro* (28, 29, 31, 38, 39), we evaluated STAT5 binding in intact, untreated male rats, and we tested the hypothesis that STAT5 binding to chromatin in liver *in vivo* is dynamically

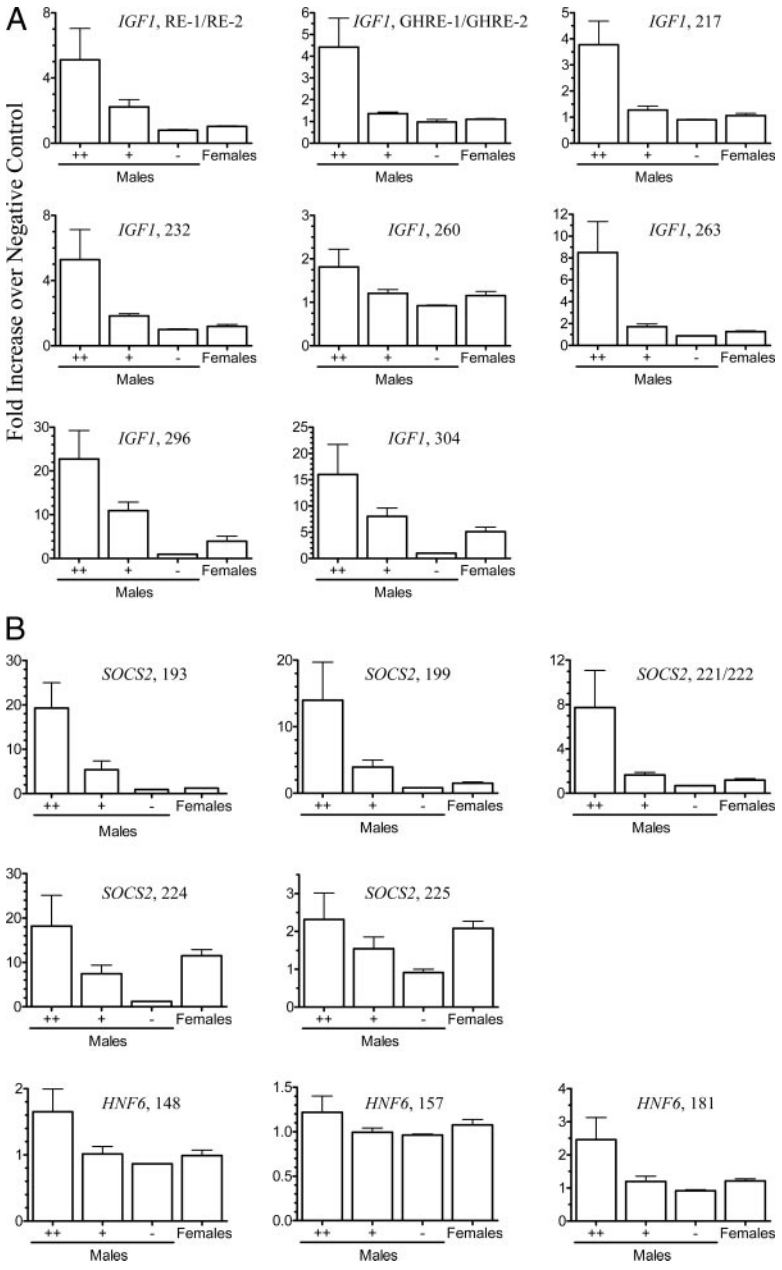

FIG. 3. ChIP analysis of STAT5 binding in female liver and in male livers that differ in STAT5 activity content. STAT5 binding to the indicated regions in *IGF-I* (*A*), *SOCS2* (*B*), and *HNF6* (*B*) were analyzed by ChIP using the indicated sets of male and female chromatin samples, prepared from individual untreated livers of adult female rats or from adult male rats with high (++), intermediate (+) and no (−) liver STAT5 activity, as indicated in Fig. 2 (20–25 μg DNA per sample). The ChIP enrichment at each STAT5 binding region was quantified and expressed as described in Fig. 1. Data shown are mean ± SEM for each group [n = 2 livers for STAT5 (−) male group and n = 4 livers for the other three groups]. STAT5 binding to *IGF-I* sites 296 and 304 and to *SOCS2* site 224 (and 225) in female liver chromatin was substantially higher than in the STAT5 (−) male samples, despite the very small difference in STAT5 activity content (*cf.* Fig. 2).

responsive to plasma GH pulsation. ChIP analysis performed with chromatin prepared from livers of male rats killed at the time of a plasma GH pulse, when liver nuclear STAT5 activity is high, enabled us to identify multiple STAT5 binding regions in each gene (Table 2). Multiple STAT5 binding regions were reported previously for *IGF-I* in mouse liver (29), but only one STAT5 region was previously identified for *SOCS2* (38) and for *HNF6* (39). Thus, these three genes, as well as several other GH-responsive genes examined here (Fig. 6), all contain multiple STAT5 binding sites. Notably, all of the STAT5 sites identified previously for *IGF-I*, *SOCS2*, and *HNF6* under nonphysiological conditions or *in vitro* were found to bind STAT5 in male liver *in vivo*.

The majority of the STAT5 binding sites identified in this study were paired sites, containing either one or two consensus STAT5 sequences (Table 2). STAT5 may bind to these sites in tandem, as tetramers, which can be expected to bind with a high overall binding affinity. *In vitro* binding studies demonstrate, however, that STAT5b, which is the major (>90%) STAT5 form in liver (50, 51), is less likely to form tetramers on DNA than STAT5a (56, 58). Nevertheless, two of the strongest binding sites for liver STAT5 (*IGF-I* site 296 and *SOCS2* site 224) contain either three (site 296) or two (site 224) STAT5 consensus sequences, raising the possibility that these sites may show enrichment for STAT5a. The tandem location of two consensus STAT5 sites served as a good predictor of STAT5 binding in the case of *IGF-I* and *SOCS2*. For *HNF6*, however, this prediction failed; its paired two-consensus sequence at site 12 (supplemental Table S3) did not bind STAT5. In the case of *RGS3*, *SPIN2A* and *SPIN2B*, which are rapidly induced by a physiological pulse of GH in rat liver (supplemental Table S8) (52), the STAT5 sites identified were all paired sites. STAT5 binding sites associated with *MUP* genes, on the other hand, were single consensus sites (Table 2).

The highest predictive score was achieved when STAT5 site predictions were based on a combination of two criteria: 1) presence of a consensus STAT5 binding sequence and 2) recognition by STAT5 matrix 459. Together, these criteria predicted all of the *IGF-I*, *SOCS2*, and *HNF6* STAT5 binding sites identified here, with the exception of nonconsensus STAT5 sites, which would necessarily be missed using this approach. Multiple STAT5 sites in *RGS3* and in various *SPIN2* and *MUP* genes were also correctly predicted. Three of the nonconsensus sites enriched by STAT5 ChIP competed for STAT5 binding *in vitro* (*HNF6* sites 148 and 157 and *IGF-I* site 232), supporting the conclusion that these sites bind STAT5 *in vivo*. Two other nonconsensus sites that gave ChIP signals did not compete for STAT5 binding *in vitro* (*SOCS2* sites 222 and 225), suggesting that the ChIP enrichment in these genomic regions reflects STAT5 binding to neighboring sites. Because the evaluation of predictors was primarily based on ChIP data obtained for *IGF-I* and *SOCS2*, the above two predictors can be best applied to genes that are similar to *IGF-I* and *SOCS2*, *i.e.* genes that respond rapidly and strongly to GH, and in a sex-independent manner. Such genes are likely to contain high-affinity STAT5 binding sites that match the consensus sequence. Genes characterized by other patterns of GH respon-

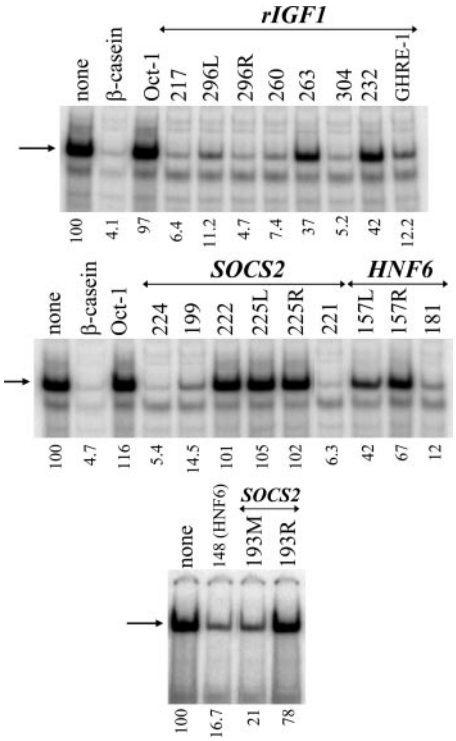

FIG. 4. *In vitro* STAT5 binding determined by competitive EMSA analysis of predicted STAT5 sites. STAT5-positive male rat liver extract was incubated with ^{32}P-labeled double-stranded STAT5 binding probe from the β-casein promoter and a 100-fold molar excess of unlabeled oligonucleotide corresponding to each of the indicated STAT5 binding sites (supplemental Table S4). In cases where the predicted STAT5 binding region (Table 2 and supplemental Table S3) was long and contained a nonbinding sequence in the middle, two oligonucleotide probes (designated L and R) were synthesized to assay STAT5 binding to sequences to the left and to the right, respectively, of the intervening sequence. Probes for site 193 represented the middle (193M) or the right (193R) portion of the full sequence containing two extended nonbinding intervening sequences. Oct-1 served as a non-STAT5 binding (control) DNA sequence. STAT5-binding regions GHRE-2, RE-1, and RE-2 of *IGF-I* were shown previously to bind to rat STAT5 *in vitro* (28, 31) and were not included in these analyses. Numbers *above* each lane indicate the STAT5 site being tested, and numbers *below* each lane represent the intensity of the STAT5-DNA band (marked with an *arrow*) relative to its intensity in the absence of an unlabeled competitor (the first lane of each gel, none, 100%).

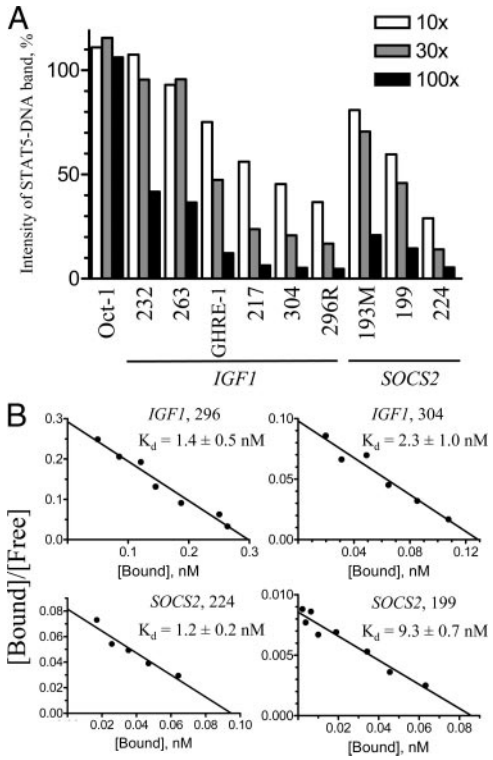

FIG. 5. Analysis of STAT5 binding affinity. A, The ability of predicted STAT5 binding sites to compete for binding to STAT5 *in vitro* was tested by EMSA as described in Fig. 4 but using 10-, 30-, and 100-fold molar excess of unlabeled competing oligonucleotides representing the indicated predicted STAT5 binding sites of *IGF-I* and *SOCS2*. Each *bar* represents the intensity of the STAT5-DNA band in the presence of the competing oligonucleotide expressed as percentage of the signal in the absence of a competing probe (set at 100%). *Numbers below the bars* indicate the STAT5 site being tested. Oct-1 is the non-STAT5 binding (control) DNA sequence. For each gene, probes are ordered based on increasing competitive binding activity. B, K_d values for binding of STAT5 to four STAT5 binding sites (sites 296 and 304 of *IGF-I* and sites 199 and 224 of *SOCS2*) were determined by Scatchard plot analysis of EMSA data. Nuclear extract prepared from GH-stimulated 293T cells transfected with rat GH receptor and rat STAT5b expression plasmids was incubated in the presence of 0.25 nM ^{32}P-labeled double-stranded oligonucleotides for sites 296, 304, 199, or 224 and increasing amounts of the same but unlabeled probe. The intensity of the STAT5-DNA band and the free probe on the gel was quantified, and the results were plotted. K_d values shown are mean ± SEM values for two to five independent determinations calculated from the negative inverse of the Scatchard plot slopes. Shown in B are representative experiments for each probe. For IGF-I site 296, a truncated oligonucleotide probe (296R) representing part of the predicted sequence for this site and containing two of the three consensus STAT5 binding sequences was used in the assay in A, whereas a full-length DNA probe containing all three consensus STAT5 binding sequences of site 296 was used for the K_d determination (B).

siveness may contain weaker, noncanonical binding sites. To predict such nonconsensus sites, a better strategy might be to first filter sites that are found in all four species and then select sites that are predicted by matrix 459 or by any of the other eight matrices but with a higher (more stringent) Possum score.

The three verified *HNF6* STAT5 binding regions identified here correspond to 14% of the sites tested, as compared with 38 and 25% of the sites tested for *IGF-I* and *SOCS2*, respectively. STAT5 binding at all three *HNF6* regions was weak in comparison with the strong STAT5 binding regions associated with *IGF-I* and *SOCS2*. Moreover, STAT5 binding at *HNF6* region 181, which gave the strongest *HNF6* STAT5 binding signal, was conserved in only three of the four species (absent in mouse), in contrast to the conservation of the strongest binding regions of *IGF-I* and *SOCS2* across all four species. Conceivably, *HNF6*, whose expression is approximately 3-fold higher in female than in male liver (40, 41), may be regulated by other, stronger STAT5 binding sites that were not identified here. These could include STAT5 sites not conserved across species, in view of the fact that site 181 was not conserved in the mouse yet was the strongest of the three *HNF6* sites identified. In addition, GH may in part regulate *HNF6* via STAT5-independent mechanisms.

Comparison of STAT5 binding in livers of male rats with different levels of STAT5 activity revealed a direct relationship between STAT5 binding to chromatin and the STAT5 activity content of the livers. Because STAT5 activity in male liver exhibits a very strong positive correlation with the occurrence of a pulse in the plasma GH profile (25, 26), these findings indicate that STAT5 binding to liver chromatin is

TABLE 3. Summary statistics for the two best individual predictors for STAT5 binding sites

Predictor(s)	TP	TN	FP	FN	Sensitivity[a]	Specificity[b]	PPV[c]	NPV[d]	Accuracy[e]
≥1 consensus site in rat	15	46	15	3	0.83	0.75	0.50	0.94	0.77
Matrix 459	17	45	16	1	0.94	0.74	0.52	0.98	0.78
Matrix 459 + consensus site	15	50	11	3	0.83	0.82	0.58	0.94	0.82

True positives (TP), true negatives (TN), false positives (FP), false negatives (FN), sensitivity, specificity, PPV, NPV, and accuracy (see definitions in Table footnotes) for the two best individual predictors and the combination of the two are listed.

[a] Sensitivity = TP/(FN + TP); [b] Specificity = TN/(FP + TN); [c] PPV = TP/(TP + FP); [d] NPV = TN/(FN + TN); [e] Accuracy = (TP + TN)/(TP + TN + FP + FN).

dynamic; *i.e.* it is directly induced by GH and cycles on and off chromatin in response to each plasma GH pulse. In males, STAT5 binding to these sites during a plasma GH pulse [STAT5 (++) livers; Figs. 2 and 3] is followed by a near-complete loss of STAT5 binding between GH pulses [STAT5 (−) livers]. Few of the STAT5 binding regions of *IGF-I* and *SOCS2* that were occupied in high STAT5 activity male liver were also occupied in female liver, and in the case of the *MUP* genes, none was detectably occupied in females. These latter sites may contribute to GH regulation of the *MUP* genes, whose expression is highly male specific and is, in part, STAT5 dependent (14, 59).

The three regions of *IGF-I* and *SOCS2* that bound STAT5 in female liver were among the strongest binding regions in males (*IGF-I* sites 296 and 304 and *SOCS2* site 224). Given the low STAT5 activity that is generally found in female rat liver (Fig. 2) (26), STAT5 activity may simply be too low for STAT5 to bind at the other sites in female liver. Indeed, *IGF-I* sites 296 and 304 and *SOCS2* site 224 are presently shown to be high-affinity binding sites (K_d ~1–2 nM), whereas *SOCS2* site 199, where STAT5 binding was observed only in males, is a low-affinity site (K_d ~9 nM). This finding indicates that sex-specific STAT5 binding, *e.g.* linked to the expression of sex-specific genes, is more likely to occur at low-affinity STAT5 sites than at high-affinity sites. This hypothesis is further supported by our finding that the male-specific STAT5 binding sites in *MUP* genes bind STAT5 with low affinity (Fig. 6D). Other factors, such as differences in chromatin structure or the presence or absence of proteins that modulate STAT5 binding to DNA, could also contribute to the sex differences in STAT5 binding seen at some sites. Given our finding that a majority of the STAT5 binding sites in *IGF-I* and *SOCS2* bound STAT5 in a male-specific manner, an important question is whether all of these sites contribute to gene transcription in liver *in vivo*, insofar as these two genes do not show sex dependence (supplemental Table S8). Conceivably, *IGF-I* and *SOCS2* transcription could be primarily regulated by STAT5 binding to sites that show high STAT5 binding in both male and female liver, *i.e.* the high-affinity sites. Although the occupancy of those sites by STAT5 in female liver is somewhat lower than in males, it is presumably more persistent due to the persistence of low-level GH signaling to STAT5 in females (27). Additional experiments will be required to address these questions.

Materials and Methods

Prediction of STAT5 binding sites

A computational phylogenetic footprinting method was developed to predict the occurrence of STAT5 binding sites that are conserved across the rat (rn4), mouse (mm8), human (hg18), and dog (canFam2) genomes for the *IGF-I*, *SOCS2*, and *HNF6* genes. Sequences encompassing 100 kb of the upstream region, the full coding sequence region, including all exons and introns, and either 3 kb (*IGF-I* and *SOCS2*) or 25 kb (*HNF6*) downstream of the coding sequence, were first scanned individually for each species for the occurrence of STAT5 binding sites, *i.e.* genomic sequences that match position weight matrices that describe a binding site for STAT5. A total of nine position weight matrices were used, five of which were obtained from the TRANSFAC database (60). One of the TRANSFAC matrices, derived from a study of STAT5 binding to synthetic oligonucleotides (56), includes sequences that contain a STAT5 consensus site as well as sequences that bind STAT5 but do not match the STAT5 consensus sequence $TTCN_3GAA$. To better detect nonconsensus STAT5 binding sites, this matrix was separated into two: one for paired STAT5 binding sites with one consensus sequence (M00AS01) and another for paired sites with no consensus sequence (M00AS02). Two other matrices were generated from a set of published STAT5 binding sequences (56) that were not represented in any of the TRANSFAC STAT5 matrices: one for paired STAT5 binding sites with a 7-bp spacer sequence (M00AS03) and another for paired STAT5 binding sites that were weak binders (M00AS04). The nine STAT5 matrices used in this study are provided in supplemental Table S1.

The matrix scanning tool Possum (61) was used to find sites in each genomic sequence that match one or more of the nine STAT5 matrices, and all hits that met a Possum threshold score of 5 were stored. Sequences from the four species examined were then aligned to identify hits that are conserved across species, which was accomplished as follows. Pairwise alignments were generated between the rat sequence and each of the other three species using two different algorithms for rapid global alignment, AVID (62) and LAGAN (63). The stand-alone version of the program VISTA (64) was then used to identify conserved regions in each pairwise alignment. Conserved regions were defined as segments at least 100 nt in length that are at least 70% identical between the two sequences. For each of the sites predicted in the rat sequence by Possum, the best pairwise alignment with each of the other three species, obtained from either AVID or LAGAN, was chosen. Sites that were shared across all four species were considered to be most likely to be functional in STAT5 binding. Each predicted STAT5 binding site was classified according to whether it contains the consensus STAT5 binding sequence $TTCN_3GAA$ in rat, which was used as the reference species. STAT5 binding sites were also classified as single or paired. A paired site was defined as one that has an overall length that is greater than twice the length of a single binding site for STAT5 (12 nt, including 3-nt flanking sequence), with a maximum intervening sequence length of 50 nt. Predicted STAT5 sites less than 50 nt apart were concatenated and classified as paired sites. The above analysis was automated using a Perl program, which is available upon request.

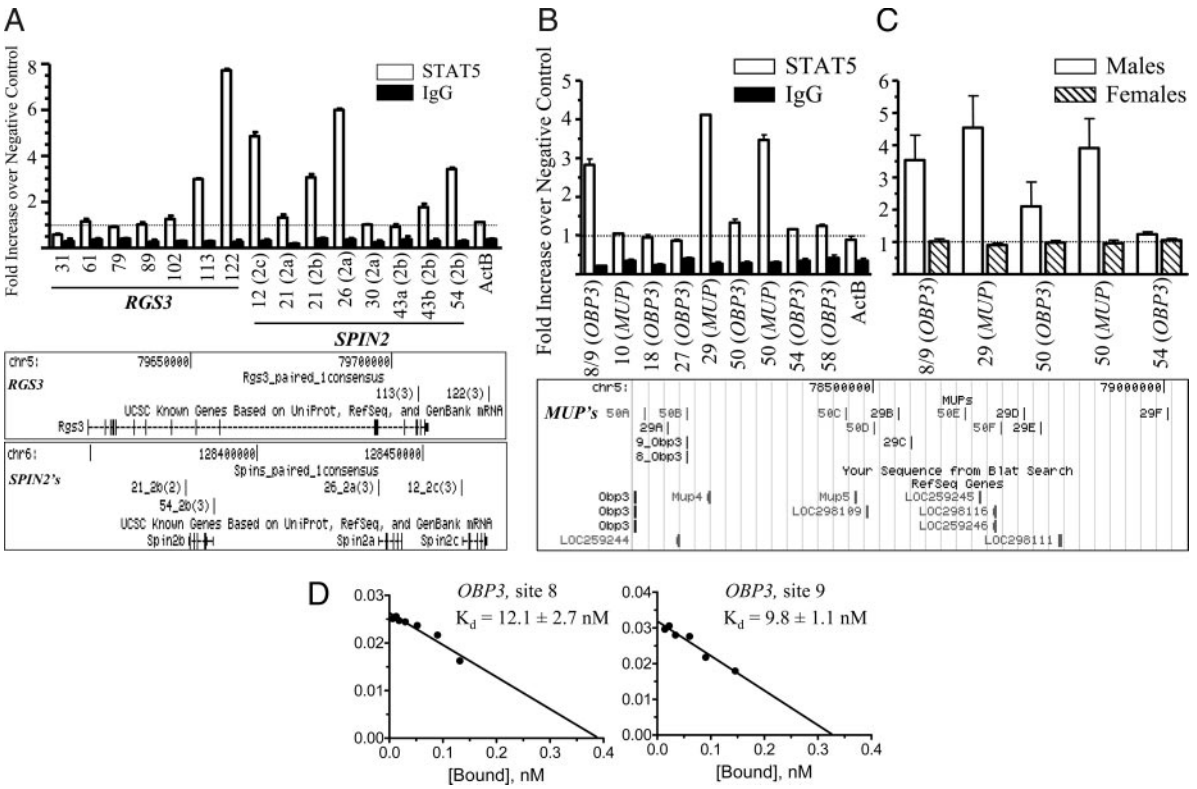

FIG. 6. STAT5-binding regions identified in *RGS3*, *SPIN2*, and *MUP* genes. Cross-linked chromatin samples prepared from frozen untreated male liver with high content of active STAT5 (sample 12 in Fig. 2) were precipitated with STAT5-antibody or normal rabbit IgG, and the abundance of STAT5 binding regions predicted in the *RGS3* and *SPIN2* (A) or *MUP* (B) loci was quantified by real-time PCR. Data were normalized and expressed as in Fig. 1 and shown as mean ± range of two determinations for each predicted STAT5 site. *Text in parentheses* in A indicates which of the three *SPIN2* genes (*SPIN2A*, *SPIN2B*, and *SPIN2C*) this site is located closest to. *Text in parentheses* in B indicates whether the STAT5 binding site was assayed for a single *MUP* gene (gene *OBP3*) or for several *MUP* genes together (*MUP*). ActB is a region within the third intron of β-actin that served as an additional negative control; this region is devoid of any predicted STAT5 binding sites, and β-actin is not regulated by GH. Genomic locations of the sites positive for STAT5 binding are shown *below* each graph, as in Fig. 1. *Below* A, *numbers in parentheses* indicate the number of species (for *RGS3* sites) or the number of genes (for *SPIN2* sites) in which the local genomic region encompassing the site is at least 70% identical. The *SPIN2* region STAT5 sites shown on the genome browser window are present in all three *SPIN2* genes, except for *SPIN2* site 12, which is found in *SPIN2B* only. *Below* B, the *MUP* region RefSeq genes shown correspond to the eight known rat *MUP* genes; these are given various *MUP*, *OBP*, and *LOC* designations, as indicated. The STAT5-positive *MUP* sites designated 29 and 50 were interrogated with generic *MUP* region primers; each amplicon amplifies genomic sequences associated with six distinct sites, numbered 29A through 29F and 50A through 50F, as shown. Sites 8 and 9 are close, but distinct STAT5 sites that were interrogated by PCR primers mapping to the specific genomic region as indicated. C, STAT5 binding to the indicated *MUP* region sites (see B) was analyzed by ChIP in the sets of male and female liver chromatin samples prepared from individual untreated adult female rats and male rats with high liver STAT5 activity (samples shown in lanes 1–3 and 11–13, respectively, in Fig. 2). The abundance of STAT5 binding regions was quantified and expressed as described in Fig. 1. Data are the mean ± SEM for each group (n = 3). *Dashed horizontal lines* indicate ChIP activity of the negative control. D, Scatchard analysis of STAT5 binding to sites 8 and 9 of the *OBP3* gene (n = 2–3 determinations, mean ± SEM, as in Fig. 5B). The sequences of these sites are identical to those of several other STAT5 binding sites predicted in various *MUP* genes, as indicated in supplemental Table S9.

Chromatin cross-linking and DNA fragmentation

Two different methods were used to prepare cross-linked chromatin from Fischer 344 rat liver. For the experiments shown in Fig. 1, a high STAT5 activity adult male rat liver sample was used to screen the predicted STAT5 binding regions of *IGF-I*, *SOCS2*, and *HNF6*. Chromatin was purified from freshly isolated liver nuclei that were immediately cross-linked with formaldehyde using a procedure adapted from a mouse liver protocol (65), followed by sonication, as detailed in supplemental Materials and Methods. For all other analyses, cross-linked samples were prepared from frozen livers excised from intact, untreated adult male and female rats (12–13 wk old). Livers were stored at −80 C and processed for cross-linking as described (31), with modifications, followed by sonication, as detailed in supplemental Materials and Methods. All animal protocols were approved by the Boston University Institutional Animal Care and Use Committee.

ChIP

All steps were performed at 4 C, unless indicated otherwise, using cross-linked samples prepared as described above. Samples were prepared from frozen liver and diluted into immunoprecipitation buffer [20 mM Tris-HCl (pH 8.1), 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium dodecyl sulfate (SDS)] containing complete protease inhibitor cocktail, or were prepared from cross-linked nuclei and diluted in RIPA buffer [50 mM Tris HCl (pH 8.1), 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS] containing complete protease inhibitor cocktail, as described in the supplemental Materials and Methods. Samples were precleared for 1 hr with Protein A Sepharose CL-4B beads (GE Healthcare, Piscataway, NJ) (50% slurry in immunoprecipitation buffer or RIPA buffer containing 1 mg/ml BSA and 200 μg/ml salmon sperm DNA: 40 μl slurry/ml of sample). Duplicate 50-μl aliquots (input) were taken from each precleared sample, and the remainder was divided into 0.9- to 1-ml aliquots. STATS antibody N-20 (sc-836 or sc-836X; Santa Cruz Biotechnology, Santa Cruz, CA) or normal rabbit IgG (sc-2027; Santa Cruz) was added to duplicate or triplicate aliquots (6 μg IgG/aliquot) followed by incubation overnight. After a 2-h incubation with protein A-Sepharose CL-4B beads (40 μl 50% slurry per aliquot), the aliquots were washed twice with immunoprecipitation or RIPA buffer (1 ml per wash per aliquot), twice with immunoprecipitation or RIPA buffer containing 0.5 M NaCl, and twice with TE buffer [10 mM Tris HCl (pH 8.1), 1 mM EDTA]. Extraction of

DNA from the beads was performed as described (66). Briefly, 100 μl 10% (wt/vol) Chelex 100 resin (Bio-Rad Laboratories, Hercules, CA) in water was added to the washed protein A-Sepharose CL-4B beads and the samples were boiled for 10 min to reverse cross-linking. After treatment with 20 μg proteinase K (Bioline, Taunton, MA) for 30–40 min at 56 C, the samples were boiled for 10 min, and supernatants (80 μl) were collected. Water (120 μl) was added to the beads, the samples were vortexed, and the supernatants were collected and pooled with the first set of supernatants to give 200 μl total supernatant volume. The samples were stored at −20 or −80 C and used undiluted in quantitative real-time PCR (qPCR) assays. Input samples (see above) were incubated for 6 h at 65 C in the presence of 0.2 M NaCl to reverse cross-linking, followed by successive treatments with 5 μg ribonuclease A (Novagen, Gibbstown, NJ) for 30 min at 37 C (not performed for samples prepared from cross-linked nuclei) and with 20 μg proteinase K for 2 h at 56 C. DNA was extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and precipitated overnight with ethanol in the presence of 10 μg glycogen (Ambion, Austin, TX) per input sample. The pellets were washed with 95% ethanol, dissolved in 50 μl water, and stored at −20 C. For qPCR, the input samples were diluted 50 times in water containing 50 μg/ml yeast RNA (Ambion).

Real-time PCR

Triplicate 5-μl real-time PCR mixtures, each containing Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA), 312 nM each qPCR primer, and 0.5–1.5 μl DNA template were loaded onto a 384-well plate and run through 40 cycles on an ABS 7900HT sequence detection system (Applied Biosystems). Primer sequences and the chromosomal positions of each amplicon are listed in supplemental Tables S2 and S3, respectively. The results for each STAT5 binding region for a given liver sample were derived from averages of duplicate or triplicate immunoprecipitation samples. Data were normalized to input and are presented as fold increase over negative control. For a negative control, we used an amplicon centered within a 2300-bp segment in the 5′ distal region of the rat *IGF-I* gene, which is devoid of any predicted STAT5 binding sites (supplemental Table S3). Similar results were obtained using a second negative control, from a STAT5 site-deficient region of the rat β-actin gene (supplemental Table S3). Data obtained in ChIP analysis carried out using normal rabbit IgG in place of STAT5 antibody N-20 corresponds to an additional control and is presented in Figs. 1 and 6 for each qPCR primer pair.

Preparation of liver homogenates and EMSA

Liver homogenates were prepared from frozen rat liver tissues as described (27) by homogenizing pieces of liver on ice in homogenization buffer [10 mM Tris-Cl (pH 7.6), 1 mM EDTA, 250 mM sucrose] containing complete protease inhibitor cocktail and PhosSTOP phosphatase inhibitor cocktail (Roche Diagnostics, Indianapolis, IN). The homogenates were centrifuged for 20 min at 9000 rpm at 4 C in a microfuge, and the supernatants were stored at −80 C. Liver homogenates were used in all EMSA experiments except for the determination of dissociation constants (see below), where nuclear extracts from transfected and GH-stimulated 293T cells were used. EMSA analysis of STAT5 binding sites was performed as described previously (27) with modifications, as detailed in supplemental Materials and Methods. Oligonucleotides used in EMSA are listed in supplemental Table S4.

Dissociation constant for STAT5 binding sites

K_d values were determined by EMSA analysis of STAT5 binding to oligonucleotides representing sites 224 and 199 of rat *SOCS2*, sites 296 and 304 of rat *IGF-I* and sites 8 and 9 of rat *OBP3* (see supplemental Table S4). These assays used nuclear extract (7.5 μg protein per EMSA reaction mixture) prepared from 293T cells transfected with expression plasmids for rat GH receptor and mouse STAT5b and stimulated with GH for 30 min. 293T nuclear extracts were prepared using a NucBuster protein extraction kit (Novagen) and were kindly provided by Dr. Rosana D. Meyer of this laboratory. A fixed amount of the ^{32}P-labeled double-stranded probe (0.25 nM) was mixed with an increasing amount of the same but unlabeled oligonucleotide, up to 100-fold molar excess. The bands corresponding to STAT5b-bound DNA and free probe were quantified using ImageQuant software, and K_d values were calculated by Scatchard plot analysis.

Evaluation of matrices and parameters used for prediction of STAT5 binding sites

A total of 79 prospective STAT5 binding sites in *IGF-I*, *SOCS2*, and *HNF6* were included in the evaluation. These 79 sites include 64 sites for which real-time PCR primer pairs were designed (primary sites) plus 15 adjacent predicted sites in the rat genome located close enough to the primary site to be indistinguishable by ChIP. The adjacent sites were selected as follows. In the case of primary sites that tested negative by ChIP, the adjacent sites were those located within a 150-bp window from each end of the amplicon. For primary sites that tested positive (*i.e.* showed ChIP enrichment), only those adjacent sites that contained a consensus STAT5 binding sequence and were identified by STAT5 matrix M00459 were included in the count. The adjacent sites also had to be located within 300 nt of the end of each amplicon in the case of weak primary binding sites (sites GHRE-1/GHRE-2, 217, 232, and 260 of *IGF-I* and sites 148, 157, and 181 of *HNF6*) or within 150 nt from the end of each amplicon in the case of strong binding sites (sites RE-1/RE-2, 263, 296, and 304 of *IGF-I* and sites 193, 199, 221/222, and 224 of *SOCS2*). Each of the 79 sites was classified according to the following 19 possible predictors of binding: presence of a paired STAT5 site; presence of at least one consensus STAT5 site in rat; presence of at least one consensus STAT5 site in any species; STAT5 site found in three other species; STAT5 site found in at least two other species; STAT5 site found in at least one other species; region of STAT5 site conserved in three other species; region of STAT5 site conserved in at least two other species; region of STAT5 site conserved in at least one other species; STAT5 site that is recognized by each of the nine STAT5 binding site matrices, with each matrix considered individually; and STAT5 site that matched any of the matrices with a possum score of at least 7 instead of the default cutoff score of 5. The 79 tested sites were classified as true positives (TP), false positives, (FP), true negatives (TN), and false negatives (FN) according to each of the 19 predictors individually, and the sensitivity [TP/(FN + TP)], specificity [TN/(FP + TN)], PPV [PPV = TP/(TP + FP)], NPV [NPV = TN/(FN + TN)], and accuracy [(TP + TN)/(TP + TN + FP + FN)] were calculated for each predictor (supplemental Table S5). The best individual predictors were determined by plotting sensitivity *vs.* (1 – specificity) in a receiver operating characteristic plot (supplemental Fig. S1). The predictors with the best combination of sensitivity and specificity, *i.e.* values closest to the (0, 1) point on the receiver operating characteristic plot, were then identified. The sensitivity, specificity, PPV, NPV, and accuracy were calculated for the combination of these predictors. To determine the validation rate by applying the best combination of conditions (consensus site and matrix 459) to sites predicted in four species, three or more species, or two or more species in *IGF-I*, *SOCS2*, and *HNF6*, STAT5-binding sites were again predicted in these genes using only matrix 459, and only consensus sites were chosen. Supplemental Table S6 shows how many of these sites were tested and how many sites tested positive.

Prediction of STAT5 binding sites in additional early GH-response genes

Matrix 459 was used to predict STAT5 binding sites in seven additional GH-responsive rat genes, *CALD1*, *GADD45G*, *NREP*, *SULT2A1*, *RGS3*, *SPIN2A*, and *SPIN2B*, and in the *MUP* gene family, using the methods described above for *IGF-I*, *SOCS2*, and *HNF6*. Sequences comprising 100 kb upstream of the transcription start site, the full coding sequence region, including all exons and introns, and 25 kb downstream of each coding sequence were scanned for matrix 459 sites, and for all genes except the *SPIN2* and *MUP* genes, rat, mouse, human, and dog sequences were aligned. In the case of the *SPIN2* genes,

sequences for *SPIN2A* (NM_012657), *SPIN2B* (NM_182474), and *SPIN2C* (NM_031531) were aligned, with *SPIN2B* taken as the reference sequence. Five *MUP* genes were analyzed: *OBP3* (NM_147215), *MUP5* (AB039828), *MUP4* (NM_198784), *LOC259245* (NM_147213), and *LOC259246* (NM_147214), with *OBP3* taken as the reference sequence. These *MUP* genes are represented by unique hybridization probes and were similarly regulated by GH in our rat microarray study (52). Sites containing a STAT5 binding consensus sequence were selected, and the sites were then classified as paired or single, and as consensus or nonconsensus sites using all nine STAT5 matrices.

Acknowledgments

Address all correspondence to: Dr. D. J. Waxman, Department of Biology, Boston University, 5 Cummington Street, Boston, Massachusetts 02215. E-mail: djw@bu.edu.

This work was supported by National Institutes of Health (NIH) Grant DK33765 (to D.J.W.). A.S. received Training Core support from the Superfund Basic Research Center at Boston University (NIH Grant 5 P42 ES07381).

Disclosure Summary: E.V.L., A.S., and D.J.W. have nothing to declare.

References

1. **Le Roith D, Bondy C, Yakar S, Liu JL, Butler A** 2001 The somatomedin hypothesis: 2001. Endocr Rev 22:53–74
2. **Edén S** 1979 Age- and sex-related differences in episodic growth hormone secretion in the rat. Endocrinology 105:555–560
3. **MacLeod JN, Pampori NA, Shapiro BH** 1991 Sex differences in the ultradian pattern of plasma growth hormone concentrations in mice. J Endocrinol 131: 395–399
4. **Pincus SM, Gevers EF, Robinson IC, van den Berg G, Roelfsema F, Hartman ML, Veldhuis JD** 1996 Females secrete growth hormone with more process irregularity than males in both humans and rats. Am J Physiol 270:E107–E115
5. **Veldhuis JD, Bowers CY** 2003 Human GH pulsatility: an ensemble property regulated by age and gender. J Endocrinol Invest 26:799–813
6. **Tannenbaum GS, Martin JB** 1976 Evidence for an endogenous ultradian rhythm governing growth hormone secretion in the rat. Endocrinology 98: 562–570
7. **Ahluwalia A, Clodfelter KH, Waxman DJ** 2004 Sexual dimorphism of rat liver gene expression: regulatory role of growth hormone revealed by deoxyribonucleic acid microarray analysis. Mol Endocrinol 18:747–760
8. **Shapiro BH, Agrawal AK, Pampori NA** 1995 Gender differences in drug metabolism regulated by growth hormone. Int J Biochem Cell Biol 27:9–20
9. **Waxman DJ, O'Connor C** 2006 Growth hormone regulation of sex-dependent liver gene expression. Mol Endocrinol 20:2613–2629
10. **Herrington J, Carter-Su C** 2001 Signaling pathways activated by the growth hormone receptor. Trends Endocrinol Metab 12:252–257
11. **Pilecka I, Whatmore A, Hooft van Huijsduijnen R, Destenaves B, Clayton P** 2007 Growth hormone signalling: sprouting links between pathways, human genetics and therapeutic options. Trends Endocrinol Metab 18:12–18
12. **Rosenfeld RG, Belgorosky A, Camacho-Hubner C, Savage MO, Wit JM, Hwa V** 2007 Defects in growth hormone receptor signaling. Trends Endocrinol Metab 18:134–141
13. **Zhu T, Goh EL, Graichen R, Ling L, Lobie PE** 2001 Signal transduction via the growth hormone receptor. Cell Signal 13:599–616
14. **Holloway MG, Cui Y, Laz EV, Hosui A, Hennighausen L, Waxman DJ** 2007 Loss of sexually dimorphic liver gene expression upon hepatocyte-specific deletion of Stat5a-Stat5b locus. Endocrinology 148:1977–1986
15. **Clodfelter KH, Holloway MG, Hodor P, Park SH, Ray WJ, Waxman DJ** 2006 Sex-dependent liver gene expression is extensive and largely dependent upon signal transducer and activator of transcription 5b (STAT5b): STAT5b-dependent activation of male genes and repression of female genes revealed by microarray analysis. Mol Endocrinol 20:1333–1351
16. **Davey HW, Park SH, Grattan DR, McLachlan MJ, Waxman DJ** 1999 STAT5b-deficient mice are growth hormone pulse-resistant. Role of STAT5b in sex-specific liver p450 expression. J Biol Chem 274:35331–35336
17. **Teglund S, McKay C, Schuetz E, van Deursen JM, Stravopodis D, Wang D, Brown M, Bodner S, Grosveld G, Ihle JN** 1998 Stat5a and Stat5b proteins have essential and nonessential, or redundant, roles in cytokine responses. Cell 93:841–850
18. **Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, Waxman DJ, Davey HW** 1997 Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proc Natl Acad Sci USA 94:7239–7244
19. **Hennighausen L, Robinson GW** 2008 Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. Genes Dev 22:711–721
20. **Ehret GB, Reichenbach P, Schindler U, Horvath CM, Fritz S, Nabholz M, Bucher P** 2001 DNA binding specificity of different STAT proteins. Comparison of in vitro specificity with natural target sites. J Biol Chem 276:6675–6688
21. **Gebert CA, Park SH, Waxman DJ** 1999 Down-regulation of liver JAK2-STAT5b signaling by the female plasma pattern of continuous growth hormone stimulation. Mol Endocrinol 13:213–227
22. **Gebert CA, Park SH, Waxman DJ** 1999 Termination of growth hormone pulse-induced STAT5b signaling. Mol Endocrinol 13:38–56
23. **Fernández L, Flores-Morales A, Lahuna O, Sliva D, Norstedt G, Haldosén LA, Mode A, Gustafsson JA** 1998 Desensitization of the growth hormone-induced Janus kinase 2 (Jak 2)/signal transducer and activator of transcription 5 (Stat5)-signaling pathway requires protein synthesis and phospholipase C. Endocrinology 139:1815–1824
24. **Waxman DJ, Ram PA, Park SH, Choi HK** 1995 Intermittent plasma growth hormone triggers tyrosine phosphorylation and nuclear translocation of a liver-expressed, Stat 5-related DNA binding protein. Proposed role as an intracellular regulator of male-specific liver gene transcription. J Biol Chem 270:13262–13270
25. **Choi HK, Waxman DJ** 2000 Plasma growth hormone pulse activation of hepatic JAK-STAT5 signaling: developmental regulation and role in male-specific liver gene expression. Endocrinology 141:3245–3255
26. **Tannenbaum GS, Choi HK, Gurd W, Waxman DJ** 2001 Temporal relationship between the sexually dimorphic spontaneous GH secretory profiles and hepatic STAT5 activity. Endocrinology 142:4599–4606
27. **Choi HK, Waxman DJ** 1999 Growth hormone, but not prolactin, maintains, low-level activation of STAT5a and STAT5b in female rat liver. Endocrinology 140:5126–5135
28. **Chia DJ, Ono M, Woelfle J, Schlesinger-Massart M, Jiang H, Rotwein P** 2006 Characterization of distinct Stat5b binding sites that mediate growth hormone-stimulated IGF-I gene transcription. J Biol Chem 281:3190–3197
29. **Eleswarapu S, Gu Z, Jiang H** 2008 Growth hormone regulation of IGF-I gene expression may be mediated by multiple distal STAT5 binding sites. Endocrinology 149:2230–2240
30. **Woelfle J, Billiard J, Rotwein P** 2003 Acute control of insulin-like growth factor-I gene transcription by growth hormone through Stat5b. J Biol Chem 278:22696–22702
31. **Woelfle J, Chia DJ, Rotwein P** 2003 Mechanisms of growth hormone (GH) action. Identification of conserved Stat5 binding sites that mediate GH-induced insulin-like growth factor-I gene activation. J Biol Chem 278:51261–51266
32. **Woelfle J, Chia DJ, Massart-Schlesinger MB, Moyano P, Rotwein P** 2005 Molecular physiology, pathology, and regulation of the growth hormone/insulin-like growth factor-I system. Pediatr Nephrol 20:295–302
33. **Yakar S, Rosen CJ, Beamer WG, Ackert-Bicknell CL, Wu Y, Liu JL, Ooi GT, Setser J, Frystyk J, Boisclair YR, LeRoith D** 2002 Circulating levels of IGF-1 directly regulate bone growth and density. J Clin Invest 110:771–781
34. **Yakar S, Liu JL, Stannard B, Butler A, Accili D, Sauer B, LeRoith D** 1999 Normal growth and development in the absence of hepatic insulin-like growth factor I. Proc Natl Acad Sci USA 96:7324–7329
35. **Wang Y, Jiang H** 2005 Identification of a distal STAT5-binding DNA region that may mediate growth hormone regulation of insulin-like growth factor-I gene expression. J Biol Chem 280:10955–10963
36. **Flores-Morales A, Greenhalgh CJ, Norstedt G, Rico-Bautista E** 2006 Negative regulation of growth hormone receptor signaling. Mol Endocrinol 20: 241–253
37. **Rico-Bautista E, Flores-Morales A, Fernández-Pérez L** 2006 Suppressor of cytokine signaling (SOCS) 2, a protein with multiple functions. Cytokine Growth Factor Rev 17:431–439
38. **Vidal OM, Merino R, Rico-Bautista E, Fernandez-Perez L, Chia DJ, Woelfle J, Ono M, Lenhard B, Norstedt G, Rotwein P, Flores-Morales A** 2007 *In vivo* transcript profiling and phylogenetic analysis identifies suppressor of cytokine signaling 2 as a direct signal transducer and activator of transcription 5b target in liver. Mol Endocrinol 21:293–311

39. **Lahuna O, Rastegar M, Maiter D, Thissen JP, Lemaigre FP, Rousseau GG** 2000 Involvement of STAT5 (signal transducer and activator of transcription 5) and HNF-4 (hepatocyte nuclear factor 4) in the transcriptional control of the hnf6 gene by growth hormone. Mol Endocrinol 14:285–294
40. **Lahuna O, Fernandez L, Karlsson H, Maiter D, Lemaigre FP, Rousseau GG, Gustafsson J, Mode A** 1997 Expression of hepatocyte nuclear factor 6 in rat liver is sex-dependent and regulated by growth hormone. Proc Natl Acad Sci USA 94:12309–12313
41. **Wiwi CA, Gupte M, Waxman DJ** 2004 Sexually dimorphic P450 gene expression in liver-specific hepatocyte nuclear factor 4α-deficient mice. Mol Endocrinol 18:1975–1987
42. **Lannoy VJ, Bürglin TR, Rousseau GG, Lemaigre FP** 1998 Isoforms of hepatocyte nuclear factor-6 differ in DNA-binding properties, contain a bifunctional homeodomain, and define the new ONECUT class of homeodomain proteins. J Biol Chem 273:13552–13562
43. **Rastegar M, Lemaigre FP, Rousseau GG** 2000 Control of gene expression by growth hormone in liver: key role of a network of transcription factors. Mol Cell Endocrinol 164:1–4
44. **Rastegar M, Szpirer C, Rousseau GG, Lemaigre FP** 1998 Hepatocyte nuclear factor 6: organization and chromosomal assignment of the rat gene and characterization of its promoter. Biochem J 334(Pt 3):565–569
45. **Delesque-Touchard N, Park SH, Waxman DJ** 2000 Synergistic action of hepatocyte nuclear factors 3 and 6 on CYP2C12 gene expression and suppression by growth hormone-activated STAT5b. Proposed model for female specific expression of CYP2C12 in adult rat liver. J Biol Chem 275:34173–34182
46. **Wiwi CA, Waxman DJ** 2005 Role of hepatocyte nuclear factors in transcriptional regulation of male-specific CYP2A2. J Biol Chem 280:3259–3268
47. **Kulkarni AB, Gubits RM, Feigelson P** 1985 Developmental and hormonal regulation of α2u-globulin gene transcription. Proc Natl Acad Sci USA 82:2579–2582
48. **Prakash A, Tompa M** 2005 Discovery of regulatory elements in vertebrates through comparative genomics. Nat Biotechnol 23:1249–1256
49. **Ureta-Vidal A, Ettwiller L, Birney E** 2003 Comparative genomics: genome-wide analysis in metazoan eukaryotes. Nat Rev Genet 4:251–262
50. **Park SH, Liu X, Hennighausen L, Davey HW, Waxman DJ** 1999 Distinctive roles of STAT5a and STAT5b in sexual dimorphism of hepatic P450 gene expression. Impact of STAT5a gene disruption. J Biol Chem 274:7421–7430
51. **Ripperger JA, Fritz S, Richter K, Hocke GM, Lottspeich F, Fey GH** 1995 Transcription factors Stat3 and Stat5b are present in rat liver nuclei late in an acute phase response and bind interleukin-6 response elements. J Biol Chem 270:29998–30006
52. **Wauthier V, Waxman DJ** 2008 Sex-specific early growth hormone response genes in rat liver. Mol Endocrinol 22:1962–1974
53. **Janeczko R, Waxman DJ, Le Blanc GA, Morville A, Adesnik M** 1990 Hormonal regulation of levels of the messenger RNA encoding hepatic P450 2c (IIC11), a constitutive male-specific form of cytochrome P450. Mol Endocrinol 4:295–303
54. **Ono M, Chia DJ, Merino-Martinez R, Flores-Morales A, Unterman TG, Rotwein P** 2007 Signal transducer and activator of transcription (Stat) 5b-mediated inhibition of insulin-like growth factor binding protein-1 gene transcription: a mechanism for repression of gene expression by growth hormone. Mol Endocrinol 21:1443–1457
55. **Nelson EA, Walker SR, Alvarez JV, Frank DA** 2004 Isolation of unique STAT5 targets by chromatin immunoprecipitation-based gene identification. J Biol Chem 279:54724–54730
56. **Soldaini E, John S, Moro S, Bollenbacher J, Schindler U, Leonard WJ** 2000 DNA binding site selection of dimeric and tetrameric Stat5 proteins reveals a large repertoire of divergent tetrameric Stat5a binding sites. Mol Cell Biol 20:389–401
57. **Basham B, Sathe M, Grein J, McClanahan T, D'Andrea A, Lees E, Rascle A** 2008 In vivo identification of novel STAT5 target genes. Nucleic Acids Res 36:3802–3818
58. **Verdier F, Rabionet R, Gouilleux F, Beisenherz-Huss C, Varlet P, Muller O, Mayeux P, Lacombe C, Gisselbrecht S, Chretien S** 1998 A sequence of the CIS gene promoter interacts preferentially with two associated STAT5A dimers: a distinct biochemical difference between STAT5A and STAT5B. Mol Cell Biol 18:5852–5860
59. **Holloway MG, Laz EV, Waxman DJ** 2006 Codependence of growth hormone-responsive, sexually dimorphic hepatic gene expression on signal transducer and activator of transcription 5b and hepatic nuclear factor 4α. Mol Endocrinol 20:647–660
60. **Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E** 2006 TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 34:D108–D110
61. **Fu Y, Frith MC, Haverty PM, Weng Z** 2004 MotifViz: an analysis and visualization tool for motif discovery. Nucleic Acids Res 32:W420–W423
62. **Bray N, Dubchak I, Pachter L** 2003 AVID: A global alignment program. Genome Res 13:97–102
63. **Brudno M, Do CB, Cooper GM, Kim MF, Davydov E, Green ED, Sidow A, Batzoglou S** 2003 LAGAN and Multi-LAGAN: efficient tools for large-scale multiple alignment of genomic DNA. Genome Res 13:721–731
64. **Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer KA, Pachter LS, Dubchak I** 2000 VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics 16:1046–1047
65. **Chaya D, Zaret KS** 2004 Sequential chromatin immunoprecipitation from animal tissues. Methods Enzymol 376:361–372
66. **Nelson JD, Denisenko O, Bomsztyk K** 2006 Protocol for the fast chromatin immunoprecipitation (ChIP) method. Nat Protoc 1:179–185